

Prof. dr hab. inż. Grzegorz Wielgosiński  
Politechnika Łódzka  
Wydział Inżynierii Procesowej i Ochrony Środowiska

## **RECENZJA**

### **pracy doktorskiej mgr inż. Mariusza Filaka pt.: „Predykcja stężeń zanieczyszczeń powietrza w perceptronowych modelach regresyjnych”**

Niniejsza recenzja została przygotowana na podstawie pisma Przewodniczącej Rady Dyscypliny Naukowej Inżynieria Środowiska, Górnictwo i Energetyka Politechniki Częstochowskiej - Pani dr hab. inż. Iwony Zawieji, prof. PCz. z dnia 12 lipca 2022 roku.

#### **1. Charakterystyka pracy**

Rozprawa doktorska mgr inż. Mariusza Filaka zatytułowana „*Predykcja stężeń zanieczyszczeń powietrza w perceptronowych modelach regresyjnych*” została wykonana na Wydziale Infrastruktury i Środowiska Politechniki Częstochowskiej. Promotorem pracy jest dr hab. Szymon Hoffman, profesor uczelni. Rozprawa ta ma typową formę pracy doktorskiej. Liczy ona wraz z załącznikami 110 stron i składa się ona z 7 głównych rozdziałów oraz bibliografii.

Ocenianą rozprawę doktorską otwiera krótki, 2-stronicowy rozdział zatytułowany „wstęp”, będący wprowadzeniem do całej rozprawy. Zawiera on wiele cennych spostrzeżeń dotyczących systemu monitoringu jakości powietrza. Kolejny rozdział, liczący jedynie 10 stron, zatytułowany „część teoretyczna” zawiera skrótowy opis istniejących aktów prawnych regulujących funkcjonowanie monitoringu powietrza w ramach Państwowego Monitoringu Środowiska, wpływ zanieczyszczeń powietrza na zdrowie człowieka, podstawowe informacje na temat sztucznych sieci neuronowych i ich zastosowania w modelowaniu jakości powietrza. Wydaje się, że ta część pracy powinna być nieco bardziej rozbudowana. W trzecim rozdziale doktorant sformułował cel pracy oraz podał główne tezy badawcze. Celem pracy było poprawienie dokładności predykcji stężeń zanieczyszczeń powietrza w neuronowych modelach regresyjnych poprzez zastosowanie wielu modeli predykcyjnych, utworzonych dla

różnych podzakresów stężeń zanieczyszczeń powietrza. Natomiast tezy badawcze zostały określone w następujących sposób:

- archiwalne dane pomiarowe mogą być wykorzystane do modelowania aktualnych stężeń zanieczyszczeń;
- do modelowania predykcyjnego można wykorzystać sztuczne sieci neuronowe, które stanowią wygodne i praktyczne narzędzie do modelowania, w porównaniu do tradycyjnych metod;
- istnieje możliwość zmniejszenia błędów modelowania regresyjnego, poprzez zastąpienie jednej sieci neuronowej przez kilka sieci, obejmujących różne sektory zakresu stężeń modelowanego zanieczyszczenia.

Rozdział czwarty ocenianej rozprawy doktorskiej został przez autora zatytułowany „część badawcza”. Do przeprowadzenia analizy wykorzystano dane z lat 2011-2016 zarejestrowane na dwóch stacjach Państwowego Monitoringu Środowiska – w Zabrze oraz w miejscowości Złoty Potok – obie w województwie śląskim. Przedmiotem analizy były zmiany jednogodzinnych stężeń ozonu ( $O_3$ ), tlenku i dwutlenku azotu ( $NO$  i  $NO_2$ ), dwutlenku siarki ( $SO_2$ ), pyłu zawieszonego  $PM_{10}$  oraz tlenku węgla ( $CO$ ) w funkcji parametrów meteorologicznych: prędkości wiatru, temperatury powietrza, natężenia promieniowania słonecznego oraz wilgotności względnej powietrza w okresie lat 2011-2016. W pierwszej części rozdziału doktorant scharakteryzował dane z obu stacji pokazując zmienność sezonową dla poszczególnych zanieczyszczeń. W dalszej części przeszedł do wyboru typu i architektury sieci neuronowej, za pomocą której zamierzał modelować przebieg zmienności i dokonywać predykcji stężeń zanieczyszczeń. Modelowanie zmienności stężeń zanieczyszczeń dla badanych stacji pomiarowych prowadzono za pomocą sztucznych sieci neuronowych o architekturze perceptronu wielowarstwowego zawierającego 10 neuronów w pojedynczej warstwie ukrytej z wykorzystaniem komercyjnego programu Statistica. Doktorant przyjął, że miarą dokładności modelu będzie średni błąd całkowity (MAE) oraz pierwiastek błędu średniokwadratowego (RMSE) których wartość minimalizował. Stworzył dwa rodzaje modeli predykcyjnych – model rosnących wartości rzeczywistych (RVS) oraz model rosnących wartości predykcyjnych (PVS). W tym celu doktorant wykonał obliczenia w kilku wariantach: dla wszystkich zebranych danych dla danego zanieczyszczenia, a następnie dla podzbiorów uzyskanych poprzez podzielenie zbioru wyjściowego (posortowanego rosnąco według stężeń zamieszczenia) na 2, 4 oraz 8 podzbiorów. W przypadku stacji monitoringowej w Zabrze było to odpowiednio: 36460 danych pomiarowych (pełny zbiór), 2 razy po 18230

danych pomiarowych (2 podzbiory), 4 razy po 9115 danych (4 podzbiory) oraz 8 razy po 4557 lub 4558 danych pomiarowych (8 podzbiorów). Dla stacji w Żółtym Potoku, przy identycznych podziałach było to odpowiednio 15536 danych, 2 razy po 7768 danych, 4 razy po 3884 danych oraz 8 razy po 1942 dane. Następnie w celu zobrazowania różnic w uzyskanych przewidywanych stężeniach sporządził 4 wykresy przedstawiające porównanie przebiegów dobowych stężeń rzeczywistych zarejestrowanych na danej stacji monitoringu powietrza oraz stężeń aproksymowanych modelami regresyjnymi. Do przedstawienia wybrał po jednym dniu z okresu wiosenno-letniego i jednym z okresu jesienno-zimowego.

Kolejny rozdział pracy, to prezentacja uzyskanych wyników oraz ich dyskusja. Liczy ona aż 53 strony i zawiera wyniki modelowania stężeń jednogodzinnych ozonu, tlenu i dwutlenku azotu, dwutlenku siarki, pyłu zawieszonego PM10 oraz tlenu węgla na obu stacjach za pomocą obu modeli tj. modelu rosnących wartości rzeczywistych oraz modelu rosnących wartości predykcyjnych. Wyniki (w postaci obliczonych błędów dopasowania – MAE i RMSE) przedstawione zostały zarówno w postaci tabelarycznej (22 tabele) jak i w postaci wykresów (46 wykresów). Widać z nich jednoznacznie, że w zdecydowanej większości przypadków błąd dopasowania malał wraz z podziałem na coraz większą ilość podzbiorów, choć czasami dla podziału na 4 podzbiory i na 8 podzbiorów był on porównywalny. Ciekawostką jest, że czasami najlepsze wyniki uzyskiwano dla podziału jedynie na 2 podzakresy.

Rozdział rozprawy 6 zatytułowany jest „Podsumowani i wnioski”. Na 4 strona zawiera on podsumowania wyników modelowania przedstawionych w poprzednim rozdziale, w wielu miejscach powielając podane wcześniej (w rozdziale 5) informacje. Zestawienie wyników wszystkich modeli w tabeli (nr 30) doskonale ilustruje uzyskane wyniki, choć szkoda, że nie ma w niej wyników modelowania na pełnym zakresie danych.

W rozdziale 7 doktorant sformułował 3 główne oraz 4 szczegółowe wnioski z wykonanej pracy. Stwierdził on, że:

- archiwalne dane pomiarowe monitoringu powietrza stanowią dobrą bazę do aproksymowania stężeń zanieczyszczeń powietrza.
- sieci neuronowe stanowią praktyczne narzędzie do modelowania predykcyjnego poziomów stężeń zanieczyszczeń powietrza.
- w porównaniu do jednej sieci neuronowej dla pełnego zakresu stężeń zanieczyszczeń powietrza, zastosowanie kilku sieci neuronowych do węższych zakresów stężeń wpływa na obniżenie wartości błędów

modelowania. Prawdliwość ta została potwierdzona dla 6 podstawowych zanieczyszczeń powietrza, monitorowanych w sposób automatyczny.

Nie ulega wątpliwości, że wnioski te są słuszne, choć wniosek pierwszy i drugi wydaje się być oczywistym. Rzeczywistą nowością tej pracy jest pomysł dzielenia danych monitoringowych na podzbiory, co pozwoliło na zmniejszenie błędu modelowania (predykcji) – tym samym wniosek trzeci stanowi istotne podsumowanie wykonanej pracy, a jego rozwinięcie znajduje się w 4 sformułowanych na koniec wnioskach szczegółowych.

Uzupełnieniem pracy jest wykaz cytowanej literatury obejmujący 134 pozycje, w tym ponad połowa to pozycje anglojęzyczne, a wiele z nich pochodzi o ostatnich 5 lat, co dowodzi znaczącej aktualności pracy i jednocześnie doskonałej znajomości literatury przedmiotu przez doktoranta.

## **2. Zakres pracy, celowość podjęcia tematu**

Oceniana rozprawa doktorska mgr inż. Mariusza Filaka zatytułowana *„Predykcja stężeń zanieczyszczeń powietrza w perceptronowych modelach regresyjnych”* dotyczy bardzo istotnego zagadnienia inżynierii środowiska - oceny stopnia zanieczyszczenia powietrza i przewidywania jego zmian na podstawie zebranych danych monitoringowych. Autor zajął się praktycznym wykorzystaniem nowego, bardzo interesującego, a zarazem wydajnego narzędzia matematycznego - sztucznych sieci neuronowych do analizy zmian stopnia zanieczyszczenia powietrza rejestrowanego na 2 stacjach monitoringowych funkcjonujących na terenie województwa śląskiego oraz przewidywania tych zmian.

Począwszy od połowy lat dziewięćdziesiątych można zaobserwować duże zainteresowanie wykorzystaniem sztucznych sieci neuronowych do modelowania złożonych procesów fizycznych, fizykochemicznych oraz chemicznych w różnych obszarach nauki. Takie próby podjęto również w zakresie inżynierii środowiska, uzyskując dobre efekty, w szczególności w obszarze ochrony powietrza i analizy danych imisyjnych (monitoringowych). Podstawową zaletą sztucznych sieci neuronowych jest możliwość odwzorowania złożonych zależności pomiędzy różnymi danymi wejściowymi a oczekiwaną odpowiedzią systemu, gdy pojawiają się problemy z tworzeniem modeli matematycznych procesu opartych o prawa zachowania masy, energii i pędu. W ostatnich 20 latach szczególnie dużo ukazało się publikacji, w których autorzy wykorzystali to narzędzie matematyczne do modelowania pól imisji oraz analizy danych monitoringowych w zakresie ochrony czystości atmosfery. Niniejsza praca doskonale wpisuje się w ten nurt i zawiera wiele elementów nowości, nie

analizowanych dotychczas w znanych pracach z tego zakresu. Można więc uznać, że jest to praca, której temat i zakres został trafnie wybrany, gdyż dotyczy on istotnego problemu inżynierii i ochrony środowiska - zarządzania jakością powietrza.

W świetle wyników pracy oraz sformułowanych na jej zakończenie wniosków, można uznać, że przyjęty w pracy program badawczy został zrealizowany, przyjęte tezy badawcze udowodnione, a cel pracy osiągnięty. Biorąc pod uwagę aktualny stan wiedzy w przedmiotowym zakresie można uznać to za oryginalny dorobek naukowy doktoranta.

### **3. Uwagi krytyczne**

Recenzowana rozprawa doktorska mgr inż. Mariusza Filaka została przygotowana w sposób przemyślany i uporządkowany. Jest napisana zwięzłym, zrozumiałym językiem, a sposób przedstawienia zarówno stanu wiedzy autora dotyczącego wykorzystania sztucznych sieci neuronowych w modelowaniu jakości powietrza jak i własnych oryginalnych osiągnięć w zakresie modelowania nie budzi zastrzeżeń. Uważna lektura pracy jednak pozwala na stwierdzenie, że w paru obszarach istnieje pewien niedosyt. I tak wydaje się, że dla uzyskania wyższej jakości pracy wskazane byłoby:

- szersze opisanie w części teoretycznej pracy zarówno problemów monitoringu jakości powietrza w Polsce jak i teorii modelowania za pomocą sztucznych sieci neuronowych - część teoretyczna pracy jest stanowczo zbyt skrótowa;
- szersze ustosunkowanie się w podsumowaniu i wnioskach do problemu podziału analizowanego okresu na podokresy i szersze przedyskutowanie wpływu sposobu podziału na podokresy na dokładność modelowania.

Szkoda także, że doktorant ograniczył się do analizy danych monitoringowych jedynie z 2 stacji PM<sub>10</sub>. Przeanalizowanie danych z jeszcze jednej, dodatkowej stacji monitoringowej podniosłoby znacznie jakość pracy.

Zdaniem recenzenta, jest to pomimo sformułowanych powyżej uwag krytycznych bardzo dobra praca wnosząca elementy nowości do naszej wiedzy o stanie zanieczyszczenia atmosfery.

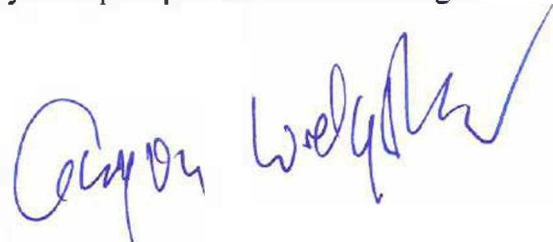
### **4. Wnioski**

Mimo sformułowanych powyżej uwag krytycznych recenzowaną rozprawę doktorską mgr inż. Mariusza Filaka należy ocenić bardzo pozytywnie. Związana jest ona z niezmiernie

ważnym obszarem inżynierii i ochrony środowiska - analizą istniejącego stanu zanieczyszczenia atmosfery oraz predykcji stanu zanieczyszczenia na podstawie danych historycznych. Praca jest spójna, przemyślana, uporządkowana i stanowi istotną nowość w zakresie ochrony czystości powietrza.

Przygotowana rozprawa dowodzi, że doktorant dobrze poznał i opanował warsztat pracy naukowo-badawczej oraz dysponuje znaczną wiedzą z zakresu reprezentowanej specjalności. Stwierdzam, że sformułowane w niniejszej recenzji uwagi mają charakter redakcyjno-dyskusyjny i nie obniżają wartości recenzowanej pracy, która spełnia wymogi stawiane rozprawom doktorskim.

Reasumując, wnoszę więc do Rady Dyscypliny Naukowej Inżynieria Środowiska, Górnictwo i Energetyka Politechniki Częstochowskiej o przyjęcie ocenianej pracy oraz dopuszczenie mgr inż. Mariusza Filaka do dalszych etapów przewodu doktorskiego.



prof. dr hab. inż. Grzegorz Wielgoński

Łódź, dn. 25 sierpnia 2022 roku